

# A Density Functional Study of the Glycine Molecule: Comparison with Post-Hartree–Fock Calculations and Experiment

D. T. NGUYEN,<sup>1</sup> A. C. SCHEINER,<sup>1</sup> J. W. ANDZELM,<sup>1</sup> S. SIROIS,<sup>2\*</sup>  
D. R. SALAHUB,<sup>2</sup> A. T. HAGLER<sup>1†</sup>

<sup>1</sup>Biosym Technologies, Inc.,<sup>‡</sup> 9685 Scranton Road, San Diego, California 92121

<sup>2</sup>Département de Chimie, Université de Montréal, CP 6128 Succursale Centre-ville, Montréal, Québec H3C 3J7, Canada

Received 14 November 1996; accepted 26 March 1997

**ABSTRACT:** The potential energy surface of un-ionized glycine has been explored with density functional theory. The performance of several nonlocal functionals has been evaluated and the results are presented in the context of available experimental information and post-Hartree–Fock quantum chemical results. The zero-point and thermal vibrational energies along with vibrational entropies play a very important role in determining the relative stability of glycine conformers; the realization of this has led to some revision and reinterpretation of the experimental results. Uncertainties in the vibrational contributions to the energy differences of several tenths of a kilocalorie/mole remain. The uncertainty in the vibrational free energy is even larger, about 1 kcal/mol. In the final analysis, we suggest that the best estimate of the electronic energy difference between the two lowest glycine conformers should be revised downward from 1.4 to 1.0 kcal/mol. Thirteen stationary points on the potential energy surface have been localized. For the majority of these, there is close agreement among various nonlocal density functionals and the post-Hartree–Fock methods. However, the second conformer (II<sub>n</sub>), which has a strong hydrogen bond between the hydroxyl hydrogen and the nitrogen of the

\* Also affiliated with the Center for Research on Computation and Its Applications (CERCA), Montréal, Québec, Canada

† Present address: Science Media, 6540 Lusk Blvd., Suite C144, San Diego, CA 92121.

‡ Now Molecular Simulations, Inc.

Correspondence to: A. T. Hagler

Contract/grant sponsors: National Institute of Technologies; NSERC; FCAR; CERCA

amine group, presents a distinct challenge. The relative energy of this conformer is extremely sensitive to the basis set, the level of correlation, or the functional used. The widely used BP86, PP86, and BP91 nonlocal functionals overestimate the strength of the hydrogen bond and predict that this conformer is the lowest energy structure. This contradicts both experiment and high-level post-Hartree–Fock studies. The adiabatic connection method (ACM) and the BLYP functional yield the correct order. The ACM method, in particular, gives energies which are in reasonable agreement with MP2, although these are somewhat low as compared with experiment. Based on this study, ACM should perform well for this type of bioorganic application, with typical errors of a few tenths of a kilocalorie/mole and only rarely exceeding 0.5 kcal/mol. © 1997 John Wiley & Sons, Inc. *J Comput Chem* **18**: 1609–1631, 1997

**Keywords:** glycine; conformational equilibrium; density functional theory; biomolecular modeling; microwave structure; adiabatic connection method

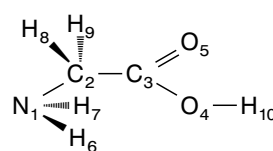
## Introduction

One of the 20 naturally occurring amino acids ( $\text{NH}_2\text{—CHR—COOH}$ ), glycine ( $\text{R=H}$ ), has been the subject of numerous detailed studies probing its un-ionized, gas-phase structure and physical properties. However, a complete and quantitative understanding of the ground state potential energy surface of glycine has proven to be a challenge to both experiment and *ab initio* theory. Beyond the intrinsic interest of the isolated glycine molecule, achieving a quantitative understanding of such energy surfaces is highly desirable in the development of accurate force fields for peptides and proteins.

The challenge derives essentially from the existence of numerous possible conformers of glycine accessed by rotating about the  $\text{N}_1\text{—C}_2$ ,  $\text{C}_2\text{—C}_3$ , and  $\text{C}_3\text{—O}_4$  bonds (see Scheme I for the atomic numbering scheme used herein). Qualitatively, the relative energetics of the various rotamers are determined by the competing effects of internal hydrogen bonding vs. steric strain and electron pair repulsions, particularly those arising from the N and O lone pairs. These are subtle energetic effects, and, consequently, as detailed in what follows, previous theoretical studies of the potential energy surface of glycine have differed in their predictions of the energetics and structures of stationary points on the surface.

## OBJECTIVES OF THIS STUDY

Over the past few years, density functional theory (DFT) has emerged as a powerful tool for the



**SCHEME I.** Atomic numbering scheme for glycine.

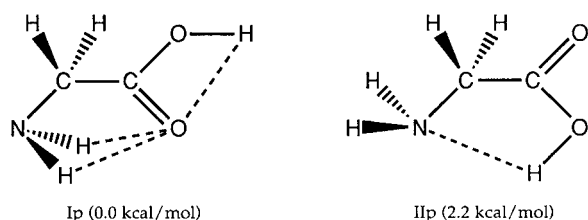
study of organic and biological molecules due, in part, to recent advances in nonlocal gradient-corrected functionals.<sup>1–3</sup> Whereas early molecular applications of DFT using functionals based on the local density approximation (LDA) were found to be inadequate for the description of hydrogen bonded and other relatively weakly bonded systems,<sup>4</sup> contemporary functionals, which include gradient corrections to LDA, have proven to yield acceptable accuracy for many types of molecular systems including those possessing hydrogen bonds.<sup>4</sup> Although DFT was developed in the 1960s by Kohn and Sham<sup>5</sup> from a solid-state physics perspective, it shows great promise for the study of large biomolecular systems because it is much faster than conventional post-Hartree–Fock (HF) methods. An example of its use in biomolecular systems is a series of studies of electron density migration in substrates on binding to the enzyme dihydrofolate reductase.<sup>6–8</sup> Due to its modest computational requirement, and its potential to be highly accurate, we are interested in applying DFT to study organic and biological systems<sup>9–12</sup> and to use the quantum mechanical potential energy surfaces in deriving molecular mechanics force fields.<sup>13,14</sup>

Before DFT can be applied routinely to peptides and, ultimately, to proteins, it must first be systematically validated. Glycine, being the simplest amino acid, is a logical (though, as we shall see,

far from trivial) starting point for such studies. The objectives of this article are: (1) to compare the "nonlocal" density approximation (NLDA) (often referred to as the GGA, or generalized gradient approximation) with high-level post-HF calculations and with experiment; (2) to evaluate the performance of a number of nonlocal functionals; (3) to investigate basis set effects on the relative energies of glycine conformers; (4) to further understand the conformational energetics of glycine and the resultant composition of the conformational equilibrium in the gas phase. To set the stage, we first review the previous work on understanding the complex potential energy surface of glycine, and some of the questions these studies raised.

## BACKGROUND

Because of its fundamental role in peptide and protein structure, as well as its small size, glycine has been the subject of extensive experimental and theoretical work. In one of the pioneering studies, in 1977, Vishveshwara and Pople<sup>15</sup> calculated single point HF/4-31G energies of glycine rotamers generated at 60° intervals about the N<sub>1</sub>—C<sub>2</sub>, C<sub>2</sub>—C<sub>3</sub>, and C<sub>3</sub>—O<sub>4</sub> bonds using fixed bond distances and bond angles. The two lowest energy structures that they found are shown in Figure 1. The lower energy structure, Ip, has a hydrogen bonded *cis*-carboxylic acid moiety and a bifurcated, long (N<sub>1</sub>—H...O5 distance of ~2.9 Å) interaction between the amine hydrogens, and the carbonyl oxygen. Iip, 2.2 kcal/mol higher in energy, has a single hydrogen bond between the hydroxyl hydrogen and the NH<sub>2</sub> nitrogen. Both of these structures have C<sub>s</sub> symmetry with a coplanar configuration of the N1—C2—C3=O5(O4—H10) atoms and with H6/H7, H8/H9 atoms being related by a plane of symmetry. Be-



**FIGURE 1.** Conformations and relative energies of the two lowest energy structures of glycine as identified from the *ab initio* study of Vishveshwara and Pople.<sup>16</sup> The notation p indicates that the heavy atoms have planar, C<sub>s</sub>, symmetry. Dashed lines represent hydrogen bonds.

cause the geometries were not energy optimized, the possibility that the actual minima were distorted from C<sub>s</sub> symmetry could not be excluded.

## EARLY MICROWAVE STUDIES OF GLYCINE DETECT ONLY CONFORMER IIP

In 1978, two independent studies of the microwave spectrum of glycine were reported.<sup>16,17</sup> To provide adequate vapor pressure of the glycine sample, the spectra were recorded at temperatures of 150° to 200°C. Although such elevated temperatures made the spectral analysis rather challenging due to the low signal-to-noise ratio and pressure broadened absorption bands, the spectral lines were analyzed and fit to a set of rotational constants. The experimentally determined rotational constants were then compared to those calculated for several possible conformers of glycine at assumed geometries. By comparing the observed and calculated rotational constants and dipole moments, both Brown et al.<sup>16</sup> and Suenram and Lovas<sup>17</sup> concluded that the source of the observed absorption lines corresponds to a glycine conformer with only one hydrogen bond, between the hydroxyl hydrogen atom and the nitrogen atom of the NH<sub>2</sub> group; that is, the higher energy *ab initio* structure Iip. Brown et al.<sup>16</sup> pointed out the "conflict" with the prediction of Vishveshwara and Pople<sup>15</sup>—the dipole moment components of Ip are predicted to be too low to be compatible with the observed spectrum. They did not detect any lines that could be attributed to Ip and, indeed, pointed out that these would be expected to be weaker, if present at all. They allowed the possibility that "the vapour contains one or more of the other conformers whose lines have so far escaped detection because they are less intense and less easily moderated by the Stark fields attainable in our spectrometer." Nevertheless, they conclude that "the most likely conformation of the glycine molecule in the vapour state is (4) [=Iip]." Suenram and Lovas,<sup>17</sup> on the other hand, stressed that "the conformer we have observed is not necessarily the only conformer present under these experimental conditions, nor does this conformer necessarily have to be the lowest energy conformation of glycine." They went on to point out that "calculations have shown that a number of conformations have low energy and several should have small dipole moments which may cause their absorption spectra to be weak and individual spectral lines may be difficult to observe by Stark-effect modulation."

## EFFECTS OF SCF GEOMETRY OPTIMIZATION ON RELATIVE ENERGY BETWEEN LOW ENERGY CONFORMERS OF GLYCINE

Because the lowest energy glycine structure predicted by theory<sup>15</sup> was not observed in the microwave experiments, Sellers and Schäfer<sup>18</sup> extended the theoretical calculations of the planar structures by optimizing the geometries of Ip and IIp at the HF/4-21G level of theory. They also calculated dipole moments of both Ip and IIp. They found no change in the relative energies of Ip and IIp due to geometry relaxation. However, they did confirm the intuition of Suenram and Lovas<sup>17</sup> and Brown et al.<sup>16</sup> that the dipole moment of IIp was significantly greater than that of Ip. The HF/4-21G-calculated dipole moment of IIp was found to be 6.5 D whereas that of Ip was 1.1 D. This results in a 35:1 ratio of spectral intensities between IIp and Ip. Thus, Sellers and Schäfer<sup>18</sup> concluded that, although conformer IIp is more visible in the microwave spectrum, conformer Ip, the global energy minimum, should be the dominant species in the gas phase and that its spectrum might be masked by the presence of IIp.

## MORE SENSITIVE MICROWAVE STUDIES CONFIRM PRESENCE OF CONFORMER IP

Because the theoretical study by Sellers and Schäfer<sup>18</sup> suggested that the microwave spectrum of Ip might be masked by the presence of IIp, Suenram and Lovas<sup>19</sup> made a number of improvements to their spectrometer to enhance the sensitivity. They reported a more sensitive microwave spectral study of gaseous glycine,<sup>19,20</sup> guided by the results of Sellers and Schäfer<sup>18</sup> and published jointly with them.<sup>20</sup> Through this they confirmed the presence of IIp and, additionally, observed a second set of spectral transitions. Although the new transitions were weak and in some cases overlapped those of conformer IIp, they were able to extract a set of rotational constants. Also, using the rate of growth technique in which the modulation of a given transition is measured as a function of the applied electric field, a rough estimate of the dipole moment was obtained which could be compared with the theoretical predictions. From this joint experimental-theoretical study<sup>20</sup> it was concluded that the new weak transitions were due to rotamer Ip, the global energy minimum.

Schäfer et al.<sup>20</sup> also predicted from calculations that a third conformer, IIIp, containing two internal hydrogen bonds lay energetically between Ip

and IIp with a relative energy of 1.9 kcal/mol (Table I). No evidence for this conformer was found in the microwave spectrum, however, and a definitive resolution of the gas phase equilibrium continued to be elusive.

Finally, the ratio of peak intensities (of the  $13_{5,8}-12_{5,7}$  transition) was used to estimate the energy difference between the  $0_{00}$  levels of the two conformers, Ip and IIp, as  $1.4 \pm 0.4$  kcal/mol (with Ip being lower in energy than IIp). The experimental uncertainty was estimated based on the sensitivity of the spectrometer and the uncertainty in the dipole moment estimate of Ip. Additionally, in the energy analysis, it was *assumed that the vibrational partition functions of the two conformers were equal*. We shall subsequently return to the effect of these assumptions and the derivation itself.

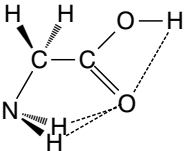
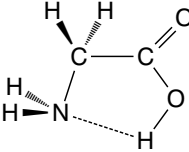
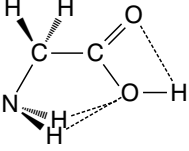
The earlier Hartree-Fock work was followed by a study in which the effects of electron correlation were calculated.<sup>21</sup> Results of this study are also tabulated in Table I. Using the geometries of Sellers and Schäfer,<sup>18</sup> Dykstra et al.<sup>21</sup> estimated that, with their self-consistent electron pair (SCEP) method (a correlated method), the energy separation between IIp and Ip decreased from 2.5 kcal/mol (at the HF level of theory) to 1.0 kcal/mol. Thus, this study demonstrated that electron correlation may significantly reduce this energy gap. Structure IIIp was not included in these studies.

## ELECTRON DIFFRACTION RESULTS

Because the microwave intensities strongly depend on the magnitude of the dipole moment of the molecule in addition to the abundance of the conformer present, Iijima et al.<sup>22</sup> felt that it was not possible to unambiguously determine the dominant glycine conformer in the gas phase from such studies. Therefore, to further characterize the low energy structure, Iijima et al.<sup>22</sup> carried out an electron diffraction study. The profile of the radial distribution function clearly showed two separate peaks for the  $N_1-O_4$  pair. The more prominent one, at 3.65 Å, indicates an  $N_1-O_4$  *anti* arrangement, and was thus assigned to conformer Ip. The secondary one at 2.8 Å shows an  $N_1-O_4$  *syn* arrangement which could either be IIp or IIIp.

From the experimental intensities, Iijima et al.<sup>22</sup> proposed that Ip is the dominant conformer (76% in the gas phase by their estimate which, however, is based on several assumptions which may be questionable). If the remaining 24% is assumed to

**TABLE I.**  
**Comparison of Microwave Experimental Results with Some of the Earlier *Ab Initio* Studies of Low-Energy Conformers of Glycine.**

Conformer	Relative energy (kcal / mol)			
	Microwave study <sup>a</sup>	HF / 4-21G <sup>b</sup>	HF / DZ <sup>c</sup>	SCEP / DZ <sup>c</sup>
 Ip	0.0	0.0	0.0	0.0
 IIp	1.4 ± 0.4	2.2	2.5	1.0
 IIIp	NO <sup>d</sup>	1.9	NO <sup>e</sup>	NO <sup>e</sup>

<sup>a</sup> Suenram and Lovas.<sup>19</sup>

<sup>b</sup> Schäfer et al.<sup>20</sup>

<sup>c</sup> Dykstra et al.<sup>21</sup> SCEP is an acronym for self-consistent electron pairs method, a correlated *ab initio* method.

<sup>d</sup> IIIp was not observed in the microwave study by Suenram and Lovas.<sup>19</sup>

<sup>e</sup> IIIp was not included in the SCEP study of Dykstra et al.<sup>21</sup>

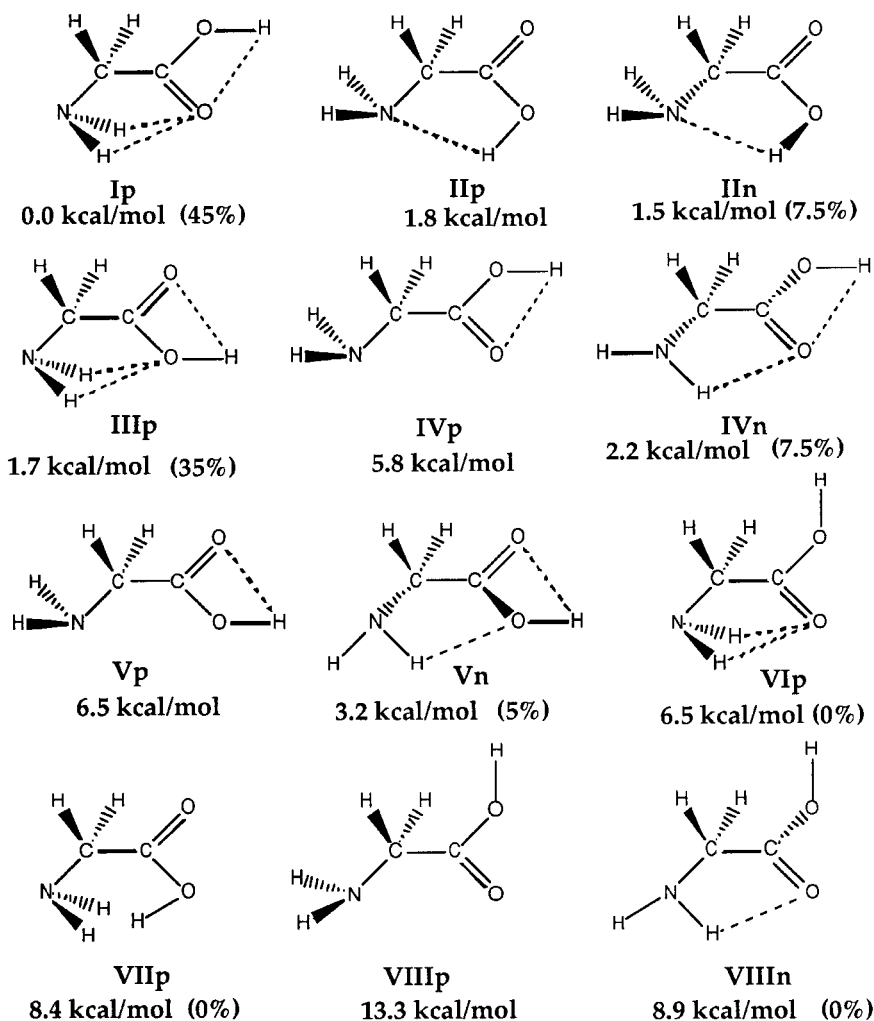
correspond to a single component, one can calculate a free energy difference of 1.3 kcal/mol (note that this was denoted as an energy in Ref. 22). Realizing that the energy difference between IIp and IIIp might be small (0.3 kcal/mol according to the HF/4-21G study of Schäfer et al.<sup>20</sup>) they repeated the analysis, assuming that IIp and IIIp are both present, which raised the value of the free energy estimate to 1.7 kcal/mol.

### EXTENSIVE SEARCH FOR STABLE GLYCINE CONFORMERS

Given the lack of conclusive experimental and theoretical results, Jensen and Gordon<sup>23</sup> carried out an extensive theoretical study of the ground state potential energy surface of glycine at the HF/6-31G\* level. In this study,<sup>23</sup> rotations about the N<sub>1</sub>—C<sub>2</sub>, C<sub>2</sub>—C<sub>3</sub>, and C<sub>3</sub>—O<sub>4</sub> bonds were systematically explored and eight unique minima and four transition states were located. Structures of these minima and transition states are shown in Figure 2. Here, we adopt the nomenclature of Császár<sup>24</sup> where the notation p denotes a planar structure with C<sub>s</sub> symmetry and n signifies a non-

planar structure. Results of the Jensen and Gordon study<sup>23</sup> are tabulated in Table II. Four of the eight minima transformed according to C<sub>s</sub> symmetry (Ip, IIIp, VIp, and VIIp), whereas the remaining four (IIp, IVp, Vp, and VIIIp) had no elements of symmetry. In addition to these eight minima, stationary points with at least one negative eigenvalue were located for conformers IIp, IVp, Vp, and VIIIp. Note that the minimum energy conformation for II was found to be nonsymmetric—IIp, rather than C<sub>s</sub>, as had been assumed in previous calculations.<sup>2,17,18</sup>

The HF/6-31G\* calculation predicts that Ip is the lowest energy structure and IIp is 3.2 kcal/mol higher in energy than Ip. This energy gap is considerably larger than the 2.2-kcal/mol value obtained from the HF/4-21G calculation of Schäfer et al.<sup>20</sup> IIp is predicted at 2.9 kcal/mol with the 6-31G\* basis set. As in the previous work of Schäfer et al.,<sup>20</sup> IIIp was calculated to be the second most stable structure, rather than the originally identified structure IIp, being 1.9 kcal/mol higher in energy than Ip at the HF/6-31G\* level of theory. Single-point MP2/6-31G\* energy calculations were carried out at each of the HF/6-31G\* stationary points



**FIGURE 2.** Stationary points on the HF / 6-31G\* glycine potential energy surface. Single-point MP2 / 6-31G\* energies along with the relative population at 500 K, as calculated from relative free energies, are given.<sup>23</sup> Here we adopt the nomenclature proposed by Császár<sup>24</sup> where the notation p indicates that the heavy atoms are planar and n signifies that these atoms are not in the same plane. The dashed lines represent hydrogen bonds.

to assess the effects of electron correlation on the relative energies of the conformers. It was found that electron correlation has further significant effects on the relative stability of the glycine conformers, as shown previously by Dykstra et al.<sup>21</sup> The MP2 single-point calculations predicted that the three lowest energy minima were Ip (0.0 kcal/mol), IIn (1.5 kcal/mol), and IIIp (1.7 kcal/mol). Thus, the relative stabilities of the glycine conformers continued to be very sensitive to theoretical method, with qualitative differences in equilibrium components resulting from higher level methods.

It is noteworthy that IIp (1.8 kcal/mol) was found to be a transition state between IIn and its mirror image rather than a minimum as assumed

in the microwave studies.<sup>19</sup> However, Jensen and Gordon<sup>23</sup> found that the zero-point energy of the minimum, IIn, was greater than the barrier height. Hence, the observed microwave structure would represent the average of the two mirror image portions of the potential energy surface.

The MP2 relative energies of the five higher energy minima were: IVn, 2.2 kcal/mol; Vn, 3.2 kcal/mol; VIp, 6.5 kcal/mol; VIIp, 8.4 kcal/mol; and VIIIIn, 8.9 kcal/mol. Although similar, they are, for the most part, lower than the HF energies. Thus, we see that the inclusion of electron correlation in general reduces the difference between the lowest energy minimum, Ip, and higher energy states, with the largest effects being in IIp and IIn where the energy drops by almost 1.5 kcal/mol.

**TABLE II.** Relative Energies (Kilocalories per Mole) and Relative Populations (%) at 500 K of Glycine Conformers as Calculated by Jensen and Gordon.<sup>23</sup>

Conformer	HF/6-31G*	MP2/6-31G* // HF/6-31G*	Relative population at 500 K (HF/6-31G*)
Ip	0.0	0.0	45%
Iip	3.2	1.8	SP <sup>a</sup>
IIn	2.9	1.5	7.5%
IIip	1.9	1.7	35%
IVp	5.6	5.8	SP <sup>a</sup>
IVn	2.2	2.2	7.5%
Vp	6.1	6.5	SP <sup>a</sup>
Vn	3.1	3.2	5%
VIp	7.0	6.5	0%
VIIp	9.3	8.4	0%
VIIIp	13.6	13.3	SP <sup>a</sup>
VIIIn	9.3	8.9	0%

<sup>a</sup> Stationary points with one or more imaginary frequencies.

Because the HF geometries were used to calculate the MP2 energies, questions remained as to the possible effect of geometry relaxation at the correlated level on the relative stability of these conformers.

### EFFECT OF TEMPERATURE ON GAS PHASE POPULATION OF GLYCINE

The effect of temperature on the relative stabilities of glycine minima was also considered by Jensen and Gordon.<sup>23</sup> They calculated the relative free energies at 500 K of the eight glycine minima and predicted the equilibrium distribution of glycine among the available minima from these values. This analysis showed that, at 500 K, the vibrational contributions caused some striking reversals in the populations that would not have been expected on the basis of the potential energy surface alone. Although IIIp was found to have a higher energy than IIn, the vibrational free energy reversed this stability dramatically, suggesting that IIIp was *five times* more abundant than Iip and almost as abundant as the minimum energy structure, Ip (35% vs. 45%). Such conclusions added further conflicting results to the glycine story as they contradicted both the microwave and electron diffraction analyses. The relatively large population of IIIp at 500 K, resulted from a single low-energy vibration of 16 cm<sup>-1</sup> for IIIp. The Jensen

and Gordon study<sup>23</sup> was also the first to suggest that other glycine conformers, IVn and Vn, may exist in the gas phase at significant levels (at 7.5% and 5%, respectively, at 500 K) and, consequently, may complicate the experimental interpretation. However, because dipole moments were not calculated, it is not possible to assess what the relative microwave intensities of these species would be expected to be.

### BASIS SET DEPENDENCE OF RELATIVE ENERGIES

At about the same time that the Jensen and Gordon<sup>23</sup> study appeared, Ramek et al.<sup>25</sup> reported an *ab initio* SCF study of five of the conformers (Ip, Iip, IIn, IIIp, and VIIIn) using 13 different small to moderate-sized basis sets ranging from STO-3G to 6-31++G\*\*. They found that, although Ip is the global energy minimum in all calculations, contradictory results are obtained for other glycine conformers. For example, Iip is a local minimum at the 4-31G, 6-31G, 6-31+G, and (10s, 5p/4s) levels, but it is a transition state with the smaller STO-3G, STO-6G, 3-21G, 4-21G, UQ10, and larger 6-31G\*\*, 6-31+G\*\*, and 6-31++G\*\* basis sets. The energy separation between IIn and Iip ranges from essentially zero to 0.44 kcal/mol. Similarly, contradictory results are obtained for conformers IIIp and VIIIn.

### EFFECTS OF GEOMETRY RELAXATION INCLUDING ELECTRON CORRELATION IN GLYCINE

Ramek et al. followed their SCF study<sup>25</sup> with an exploration of the energetic and structural effects of relaxing the geometries of Ip, IIn, and Iip, including electron correlation. In this investigation,<sup>26</sup> full MP2 geometry optimizations of these three conformers were carried out using the 6-311G\*\* basis set. Significant changes in relative energetics and structural parameters due to relaxation of the geometries were observed (summarized in Table III). Specifically, it was found that the energy difference between Iip and Ip decreased from 1.3 kcal/mol with the geometries optimized at the SCF level of theory to only 0.8 kcal/mol when geometry relaxation was allowed at the MP2 level (Table III). The MP2/6-311G\*\* energy ordering was determined to be Ip (0.0 kcal/mol), IIn (0.69 kcal/mol), and Iip (0.81 kcal/mol). Thus, based on this work, the "best" theoretical prediction for the energy difference between Ip and IIn and Iip was

**TABLE III.** Including Electron Correlation Drastically Reduces Energy Difference Between Conformers I and II of Glycine.<sup>26</sup> Geometry Optimization at Correlated, MP2 Level of Theory Leads to Further Decrease. Relative Energies in Kilocalories / Mole.

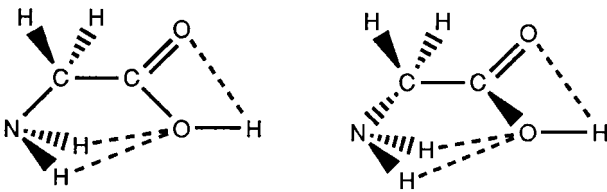
Conformer	6-311G**		
	HF (opt) <sup>a</sup>	MP2 (fixed) <sup>b</sup>	MP2 (opt) <sup>c</sup>
I <sub>p</sub>	0.0	0.0	0.0
II <sub>p</sub>	3.2	1.3	0.8
II <sub>n</sub>	3.0	1.1	0.7

<sup>a</sup> Full geometry optimization at the HF level of theory.  
<sup>b</sup> Single-point MP2 calculation using HF-optimized geometry.  
<sup>c</sup> Full MP2 geometry optimization.

smaller than the 3.2 kcal/mol at the HF level and the 1.4 kcal/mol result suggested by the microwave spectral analysis.<sup>19</sup> There was no evidence, however, to suggest that the MP2/6-311G\*\* results had converged sufficiently with respect to either basis set or electron correlation treatment.

Two recent investigations of glycine have extended the *ab initio* theory still further to explore the effects of additional basis set improvements and more complete treatments of electron correlation. In one of the studies, Császár<sup>24</sup> optimized the geometries of the 12 glycine conformers studied by Jensen and Gordon<sup>23</sup> (shown in Fig. 2) at the MP2/6-311 + + G\*\* level of theory. In addition, he identified an additional conformer, III<sub>n</sub>, which is the nonplanar equivalent of III<sub>p</sub>. Both III<sub>p</sub> and III<sub>n</sub> are shown in Figure 3. Single-point energy calculations were also carried out using an extended 13s8p3d2f basis set (termed B2). A summary of this study is presented in Table IV.

Császár found that, although conformer I<sub>p</sub> remains the lowest energy structure, its relative stability is decreased dramatically, now being only 0.51 kcal/mol more stable than II<sub>n</sub>. The planar structure II<sub>p</sub> is 0.57 kcal/mol above I<sub>p</sub>. Structure III<sub>n</sub> is only 1.45 kcal/mol higher in energy in these



**FIGURE 3.** Structures of III<sub>p</sub> and III<sub>n</sub> studied by Császár.<sup>24</sup>

**TABLE IV.** Effects of Basis Sets on Relative Energies of Glycine Conformers (kcal / mol) Calculated Using Post-Hartree–Fock Methods.

Conformer	MP2		
	6-311G* <sup>a</sup>	6-311 + + G** <sup>b</sup>	B2 <sup>c</sup>
I <sub>p</sub>	0.00	0.00	0.00
II <sub>p</sub>	0.80	0.57	0.44
II <sub>n</sub>	0.70	0.51	0.38
III <sub>p</sub>		1.59	1.67
III <sub>n</sub>		1.45	1.67
IV <sub>p</sub>		4.68	4.54
IV <sub>n</sub>		1.27	1.18
V <sub>p</sub>		5.38	5.05
V <sub>n</sub>		2.51	2.54
VI <sub>p</sub>		5.71	5.03
VII <sub>p</sub>		6.95	5.82
VIII <sub>p</sub>		11.20	10.20
VIII <sub>n</sub>		7.09	6.22

<sup>a</sup> Frey et al.<sup>26</sup> Full geometry optimization at the MP2/6-311G\* level of theory.  
<sup>b</sup> Császár.<sup>24</sup> Full geometry optimization at the MP2/6-311 + + G\*\* level of theory.  
<sup>c</sup> Császár.<sup>24</sup> Single-point MP2 energy calculation using the MP2/6-311 + + G\*\*-optimized geometries and the extended basis set contained 13s8p3d2f functions.

MP2/6-311 + + G\*\* calculations. III<sub>p</sub> is at 1.59 kcal/mol. Thus, the trend found by Frey et al.,<sup>26</sup> where inclusion of correlation and subsequent geometry optimization strongly stabilize structures II and III relative to I<sub>p</sub>, is also observed by Császár.<sup>24</sup> Single-point MP2 energy calculations using the extended basis set show that the energy gap between I<sub>p</sub> and II<sub>p</sub> and II<sub>n</sub> is further reduced with the larger basis set. II<sub>n</sub> and II<sub>p</sub> are now only 0.38 and 0.44 kcal/mol, above I<sub>p</sub>. Again, we see that the larger basis set stabilizes II<sub>p</sub> and II<sub>n</sub> relative to I<sub>p</sub>. However, conformers III<sub>n</sub> and III<sub>p</sub> are slightly destabilized.

In addition to these optimizations, Császár<sup>24</sup> also carried out single-point energy calculations using the still higher level MP3, MP4, MP $\infty$  (by extrapolation), CCSD and CCSD(T) methods with the 6-311 + + G\*\* basis set. Results from the MP2, MP4, MP $\infty$ , and CCSD(T) calculations are tabulated in Table V. There is only a small change in the glycine relative energies as the result of including a more complete treatment of electron correlation. The single-point MP4, MP $\infty$ , and CCSD(T) calculations predict slightly higher energy differences between I<sub>p</sub> and II<sub>p</sub>, but still less than 1 kcal/mol. As for conformer III<sub>p</sub>, MP2, MP4, MP $\infty$ ,



**TABLE V.**  
**Effects of Electron Correlation on Relative Energies**  
**(Kilocalories per Mole) of Glycine. Calculations**  
**Carried Out Using 6-311 + + G\*\* Basis Set at**  
**MP2 Geometry.**

Conformer	MP2 <sup>a</sup>	MP4 <sup>b</sup>	MP <sub>∞</sub> <sup>b</sup>	CCSD(T) <sup>b</sup>
Ip	0.00	0.00	0.00	0.00
Iip	0.57	0.66	0.70	0.86
IIn	0.51	1.01	0.64	
IIIp	1.59	1.54	1.52	1.63
IIIn	1.45	1.39	1.37	
IVp	4.68	4.66	4.66	
IVn	1.27	1.31	1.31	
Vp	5.38	5.34	5.34	
Vn	2.51	2.18	2.17	
VIp	5.71	5.49	5.48	5.54
VIIp	6.95	6.92	6.90	
VIIIp	11.20	10.92	10.90	
VIIIn	7.09	6.92	6.91	

<sup>a</sup> Császár.<sup>24</sup> Full geometry optimization at the MP2/6-311 + + G\*\* level of theory.

<sup>b</sup> Császár.<sup>24</sup> Single-point MPn and coupled cluster energy calculations using the MP2/6-311 + + G\*\*-optimized geometries.

and CCSD(T) predict that IIIp is between 1.52 and 1.63 kcal/mol above Ip. For the remaining glycine conformers, the relative energies vary by less than 0.4 kcal/mol with the different treatments of electron correlation.

In yet another attempt to understand the glycine potential energy surface, Hu et al.<sup>27</sup> reported fully optimized geometries of selected glycine conformers (Ip, IIn and Iip, IIIn and IIIp, Vn and Vp) at the CISD/DZP level of theory, as opposed to the MP2 optimizations of the Császár study.<sup>24</sup> Results similar to Császár's CCSD(T) calculations<sup>24</sup> were obtained.

## RECENT MICROWAVE STUDY

Godfrey and Brown<sup>28,29</sup> recently presented a study of glycine using a Stark-modulated free-expansion jet microwave spectrometer to address some of the unresolved issues experimentally. In that study, high-level *ab initio* calculations predicted that IIn is lower in energy than Iip and that other glycine conformers, in addition to those previously suggested by experiment, may also be present in the gas phase. Although Godfrey and Brown<sup>28,29</sup> detected only Ip and Iip, they could not rule out the presence of other conformers, such as IIIp, and they propose that interconversion of

species in the seeded supersonic jets could account for the "missing conformers." In what follows we shall provide estimates of relative abundance drawn from both density functional and post-Hartree-Fock theories, which are pertinent to the question.

## SUMMARY OF "GLYCINE STORY"

To summarize the "glycine story" to this point, the Ip-Iip (and IIn) energy separation is sensitive to the treatment of electron correlation and to the choice of basis set. In addition, geometry optimization was found to be critical. HF/6-311G\*\* predicted that Iip is about 3.0 kcal/mol higher in energy than Ip. Single-point MP2/6-311G\*\* energy calculations reduced the Iip-Ip energy difference to slightly more than 1 kcal/mol. Geometry optimization carried out at the MP2/6-311G\*\* level of theory further narrowed this energy gap to about 0.8 kcal/mol. With the extended basis set containing 13s8p3d2f functions, Iip was only about 0.4 kcal/mol higher than Ip, also at the MP2 level of theory. However, the Iip-Ip energy difference widened with improved treatment of electron correlation effects: Iip was calculated to be 0.66, 0.70, and 0.86 kcal/mol higher at the MP4, MP<sub>∞</sub>, and CCSD(T) level of theory, respectively. Nevertheless, because these were single-point energy calculations, the effects of relaxing the geometries are not known. Taken together, the trends with improved basis sets and higher level of correlation pointed consistently to a value for the difference in energy between Ip and Iip of around 1 kcal/mol or less, somewhat less than the value of 1.4 kcal/mol suggested from the microwave analysis.<sup>19</sup>

Furthermore, all post-HF methods predicted that IIIp was a low energy conformation which was likely to be present in the gas phase, although it was somewhat higher in energy than Iip and IIn. IIIp was estimated by theory to be about 1.5 kcal/mol higher in energy than Ip. Furthermore, vibrational free energy effects were found to increase its relative stability to the point where it comprises 35% of the gas phase population. The post-HF results contradicted the earlier HF results which found IIIp to be lower in energy than Iip (and IIn); however, the free energy was not evaluated.

Theory also suggested the presence of other low-lying glycine conformers, namely IVn and Vn at 1.27 and 2.21 kcal/mol above Ip (at the MP2/6-311 + + G\*\* level of theory). These structures were

postulated to be populated in the gas phase at 500 K. Therefore, they may complicate the interpretation of both the microwave<sup>19</sup> and electron diffraction spectra.<sup>22</sup>

However, even considering the most rigorous calculations performed to date, at the high level of accuracy required to situate these low-lying conformers with respect to each other, complete convergence has not been obtained in the energies with respect to the level of electron correlation used and the basis set. Ideally, one would carry out full geometry optimizations in the relevant regions of the potential energy surface using increasingly accurate treatments of electron correlation and increasingly larger basis sets until sufficient convergence is reached. However, the cost of traditional *ab initio* methods including electron correlation (many-electron perturbation theory, configuration interaction, coupled cluster) has limited this approach.

Thus, as we have seen, although the glycine molecule has been studied extensively by both experiment and theory, there are still some important unanswered questions regarding its potential energy surface. What conformers are actually present in the gas phase equilibrium of glycine? What are the energy differences between various glycine conformers? Why is the relative energy of IIp or IIn so sensitive to the details of the calculations? What is the nature of the IIp and IIp stationary points? What are the effects of vibrational free energy on the populations of the glycine conformers? We will go on to address these questions by supplementing the extensive calculations carried out to date with results of an alternative *ab initio* method, DFT (another DFT study,<sup>41,42</sup> which appeared while our work was in progress, is discussed later). We will compare the DFT results with those of other methods and with experiment and emphasize the common picture that emerges from the different theoretical and experimental methodologies applied to the conformational equilibrium of glycine.

The remainder of this article is organized as follows. In the next section we describe the DFT methods used in this study. We then present the results of the DFT calculations and their relation to previous work. We examine several issues including: (1) comparison of DFT with post-HF calculations; (2) assessment of the performance of a number of nonlocal functionals; (3) investigation of basis set effects on the glycine relative energies determined by DFT; (4) investigation of the correlation between hydrogen bond interactions and the

stability of glycine conformers; and (5) assessment of the composition of the conformational equilibrium of glycine in the gas phase from the picture that emerges from the diverse experimental and theoretical data. A reanalysis of the microwave results is also presented. The final section provides results and conclusions of the study.

---

## Methods

### DENSITY FUNCTIONAL THEORY CALCULATIONS

The DFT calculation described in this article were carried out using the deMon program (version 1.0) developed by St-Amant and Salahub at the Université de Montréal<sup>30</sup> and the Turbo DFT program.<sup>31</sup> All calculations were performed on Silicon Graphics Indigo R4000 and IBM RS/6000 workstations.

### NONLOCAL FUNCTIONALS AND BASIS SETS

To investigate the effects of various recent DFT functionals on both the energy and equilibrium structures of glycine, geometry optimizations were carried out using the Perdew (PP86),<sup>32</sup> Becke–Perdew (BP86),<sup>1</sup> and (BP91),<sup>33</sup> and Becke–Lee–Yang–Parr (BLYP)<sup>2</sup> nonlocal functionals, and the adiabatic connection method (ACM).<sup>34</sup>

To further evaluate basis set effects in DFT calculations and compare their importance with those found in post-HF calculations, we studied the DZVD, DZVP, and TZVP basis sets. Both the DZVD and DZVP basis sets<sup>35</sup> have a (621/41/1) contraction pattern for the first row atoms. The DZVD hydrogen atom has a (41) contraction pattern and the DZVP basis set has an additional polarization function on the hydrogen atoms. The DZVD and DZVP sets correspond to standard 6-31G\* and 6-31G\*\* basis sets in quantum chemistry, but they have been optimized specifically for DFT. The triple-zeta valence plus polarization (TZVP) basis set<sup>35</sup> has a (7111/411/1) contraction pattern for the first row atoms and a (311/1) contraction pattern for hydrogen atoms. This basis set is equivalent in size to the 6-311G\*\* basis set.

### DETAILS FOR deMON CALCULATIONS

In the deMon program, the charge density was fit analytically, whereas the exchange-correlation potentials were fit numerically on a grid. We used

the fine grid option in deMon which corresponds to a grid comprising 32 radial shells and 26 angular points per randomly rotated shell, giving rise to a total of 832 grid points/atom. At the end of each SCF procedure, the exchange-correlation contributions to the energy and energy gradients were evaluated using an augmented numerical grid containing 2968 points/atom. Geometries were optimized until the norm of the gradient was less than 0.0005 a.u. using the Broyden–Fletcher–Goldfarb–Shanno algorithm.<sup>36</sup>

## DETAILS FOR TURBO DFT CALCULATIONS

The Turbo DFT program implements a hybrid of the HF and DFT methods. It is possible<sup>37</sup> to view the Kohn–Sham equation<sup>5</sup> as strictly analogous to the standard HF procedure, except that the HF exchange potential is replaced by a local (or nonlocal) DFT exchange-correlation potential. Thus, the Fock matrix ( $F$ ) can be expressed as:

$$F = H + J + K^{XC} \quad (1)$$

where  $H$  is the one-electron Hamiltonian matrix,  $J$  is the usual Coulomb matrix, and  $K^{XC}$  is the DFT exchange-correlation matrix (which replaces the usual exchange matrix,  $K$ , in HF theory). In Turbo DFT,  $H$  and  $J$  are calculated in the same fashion as for a normal HF calculation, whereas the elements of the  $K^{XC}$  matrix are calculated by numerical integration over an atom-centered grid. Here, we use an adaptive grid with around 3000 points per hydrogen atom and more than 4000 points for first row atoms.<sup>38</sup> DFT gradients are calculated in full including both functional and weight derivatives and show exact translational invariance.

The Turbo DFT calculations were carried out using the BP91<sup>33</sup> and BLYP<sup>2</sup> exchange-correlation nonlocal gradient corrected functionals and the adiabatic connection method (ACM) proposed by Becke.<sup>34</sup> In ACM, the exchange-correlation energy ( $E_{XC}$ ) is expressed as:

$$E_{XC} = E_{XC}^{LDA} + 0.20(E_X^{exact} - E_X^{LDA}) + 0.78E_X^{B88} + 0.82E_C^{P91} \quad (2)$$

where  $E_{XC}^{LDA}$  is the Vosko–Wilk–Nusair (VWN) local exchange-correlation energy,<sup>39</sup>  $E_X^{exact}$  is the HF exact exchange energy,  $E_X^{B88}$  is the Becke88 gradient correction to the exchange energy<sup>40</sup> and  $E_C^{P91}$  is the Perdew91 gradient-corrected correlation energy.<sup>1</sup> Geometries were optimized until the maximum gradient component was less than 0.0003

a.u. and the energy change between minimization steps was less than  $10^{-6}$  hartrees.

## VIBRATIONAL FREQUENCY CALCULATIONS

To evaluate the zero-point energy (ZPE), thermal vibrational energy, and vibrational entropy contributions to the relative stability of glycine conformers, the vibrational frequencies were calculated by finite difference of the gradients at the ACM/DZVD levels of theory for all 13 glycine conformers. Furthermore, to evaluate the effects of using different functionals, we also calculated the frequencies for Ip, IIp, IIn, IIIp, and IIIn using the BLYP functional. Because the MP2 frequencies for IIIp are not available, we also evaluated these frequencies at the MP2/DZP level of theory.

## Results and Discussion

In this section, we first examine the effects of varying the basis set on the relative DFT energies of a few selected glycine conformers. We then assess the performance of various nonlocal DFT functionals. The most accurate DFT energies and free energies for the known glycine structures are then compared to both experiment and post-HF studies.

### BASIS SET EFFECTS, IN DFT, ON ENERGETICS OF GLYCINE

Before carrying out extensive DFT calculations, we first investigated effects of varying the basis set on the relative energies of glycine conformers. For this purpose, we selected a set of five representative conformers (Ip, IIp, IIIp, Vn, and VIIIIn), which include low, medium, and high energy structures. These were optimized using the BLYP<sup>2</sup> nonlocal functional with the DZVD, DZVP, and TZVP basis sets.

The relative energies of glycine conformers calculated with the three different basis sets are presented in Table VI. As noted in the Introduction, for MP2 calculations, the larger basis set tends to stabilize IIp relative to Ip. We observe a similar trend with BLYP. Relative to Ip, IIp is predicted to be at 0.33 kcal/mol with the DZVD basis set, at 0.17 kcal/mol with the DZVP set, and finally, with the TZVP basis set, the two conformers are essentially isoenergetic.

As seen from Table VI, this behavior varies with conformer. The larger basis set slightly destabilizes

**TABLE VI.**  
**Basis Set Effects on Relative Stability (Kilocalories per Mole) of Glycine Conformers Calculated Using BLYP Nonlocal Functional.**

Conformation <sup>a</sup>	BLYP		
	DZVD	DZVP	TZVP
Ip	0.00	0.00	0.00
Iip	0.33	0.17	0.02
IIip	1.34	1.44	1.49
Vn	3.01	3.01	2.72
VIIIIn	7.42	6.96	6.69

<sup>a</sup> See Figure 2 for the structures of these conformers.

IIip relative to Ip. IIip is predicted to be 1.34 kcal/mol higher than Ip at the BLYP/DZVD level of theory. Adding an extra polarization function to the hydrogen atom (DZVP basis set) increases this energy difference to 1.44 kcal/mol and going to the TZVP level yields an energy of 1.49 kcal/mol higher than Ip. At the BLYP/DZVD and BLYP/DZVP levels of theory, Vn is predicted to be 3.01 kcal/mol higher in energy than Ip. This energy gap decreases slightly to 2.72 kcal/mol when the TZVP basis set is used. The largest change in relative energy resulting from basis set changes is observed for conformer VIIIIn: the relative energy between VIIIIn and Ip decreases from 7.42 kcal/mol with the DZVD basis set to 6.96 kcal/mol with DZVP and to 6.69 kcal/mol with TZVP. These variations are similar to those observed for the MP2 method (Table IV), namely, the energies of IIip and Vp increase with the size of the basis set relative to Ip and the energy separation between VIIIIn and Ip decreases with the larger basis set. Because the relative energies of glycine conformers within 2 kcal of the global minimum

calculated with the DZVP basis set are within 0.15 kcal/mol of those obtained with the TZVP basis set and the maximum difference is 0.3 kcal/mol, and because the DFT calculations calculated using TZVP are twice as time consuming as those using DZVP, we employed the DZVP basis set for subsequent studies.

PERFORMANCE OF VARIOUS NONLOCAL GRADIENT-CORRECTED FUNCTIONALS

Another important factor in the use of DFT to simulate organic and biological systems is the performance of various nonlocal functionals. We optimized the geometries of the five glycine conformers listed in Table VI using the PP86, BP86, BP91, and BLYP nonlocal functionals, as well as the adiabatic connection method (ACM), all with the DZVP basis set. In Table VII, we tabulate the results and compare them with those of the MP2/6-311 + + G\*\* study by Császár.<sup>24</sup>

We first note that, with the exception of conformer Iip, the relative energies calculated with these five DFT methods are similar. The largest difference (0.84 kcal/mol between BP86 and BLYP) was found for conformer Iip. For conformers other than Iip, the DFT relative energies are within about 0.6 kcal/mol of the MP2 results and reflect the trends found in MP2. For Iip, discrepancies from the MP2 results of about 1 kcal/mol are found. Iip is predicted to be 0.57 kcal/mol higher than Ip by MP2. The BP86 nonlocal functional places Iip 0.67 kcal/mol *lower* than Ip. The PP86 and BP91 nonlocal functionals also give Iip as the global minimum at 0.25 and 0.49 kcal/mol, respectively, *lower* in energy than Ip. This trend is reversed with the BLYP and ACM methods which predict that Iip is slightly higher in energy than Ip, at 0.17 and 0.11 kcal/mol, respectively. Of the five DFT variants

**TABLE VII.**  
**Comparison of Nonlocal Functionals in Predicting the Relative Energies (Kilocalories per Mole) of Glycine Conformers. DFT Calculations Carried Out Using DZVP Basis Set.**

Conformer	MP2 <sup>a</sup>	BP86	PP86	BP91	BLYP	ACM
Ip	0.00	0.00	0.00	0.00	0.00	0.00
Iip	0.57	−0.67	−0.25	−0.49	0.17	0.11
IIip	1.59	1.59	1.76	1.46	1.44	1.56
Vn	2.51	2.95	3.01	3.11	3.01	3.04
VIIIIn	7.09	6.98	6.92	6.86	6.96	7.13

<sup>a</sup> Császár.<sup>25</sup> MP2 calculations carried out using the 6-311 + + G\*\* basis set.

used here, the BLYP and ACM nonlocal functionals predict the correct ordering of the glycine conformers Ip and IIp and, in general, reliably reproduce the energy ordering found by MP2. In addition, previous studies indicate that these are two of the most accurate functionals available, in agreement with these results. Thus, for the remainder of this study, we will concentrate on the BLYP and ACM results.

### COMPARISON OF DFT RELATIVE ENERGIES OF GLYCINE CONFORMERS WITH POST-HF RESULTS

To further validate the use of the BLYP nonlocal functional and ACM to model simple amino acids, we calculated the relative energies of all 13 known glycine conformers using the DZVP basis set. The DFT results are compared with the MP2 study of Császár<sup>24</sup> in Table VIII. Akin to what was seen in Table VII, the relative energies calculated with BLYP and ACM are similar to each other, within 0.4 kcal/mol, except for VIIIp, where a difference of 1 kcal/mol is found. The rms deviation between the ACM and BLYP relative energies is 0.3 kcal/mol. We also note that both the ACM and BLYP relative energies are in line with the MP2 results. The RMS deviation between ACM and MP2 and between BLYP and MP2 are both about 0.4 kcal/mol. The agreement is within 0.5 kcal/mol except for the higher energy conformers, VIIp and VIIIp, which show differences up to 0.9 kcal/mol. It is interesting to note that both BLYP and ACM confirm that conformer IVn is another low energy structure with relative energies of only 1.72 and 1.59 kcal/mol. IVn is predicted by MP2/6-311++G\*\*<sup>24</sup> to be 1.27 kcal/mol above Ip and is even lower in energy than IIIp. If IIIp is assumed to be present in the gas phase by the electron diffraction study by Iijima et al.<sup>22</sup> then IVn should also be considered. Thus, we observe that both BLYP and ACM give results for the glycine conformational energy surface which are in overall agreement with the more traditional MP2 method.

In a similar study, Lelj et al.<sup>41,42</sup> also applied DFT to study the gas phase potential energy surface of glycine. They optimized a subset of the glycine conformers studied here using the VWN local functional<sup>39</sup> and the BLYP and B3LYP nonlocal functionals using a smaller Hartree–Fock-based DZP basis set. (The B3LYP nonlocal functional is a variant of the ACM method. It uses a Lee–Yang–Parr nonlocal correlation functional,<sup>2</sup>

**TABLE VIII.**  
**Comparison with MP2 Results of Relative Energies (Kilocalories per Mole) of Glycine Conformers Calculated with DFT.**

Conformer	MP2 <sup>a</sup>	BLYP	ACM
Ip	0.00	0.00	0.00
IIp	0.57	0.17	0.11
IIIn	0.51	0.12	0.08
IIIp	1.59	1.44	1.56
IIIn	1.45	1.39	1.59
IVp	4.68	5.08	5.18
IVn	1.27	1.72	1.59
Vp	5.38	5.27	5.65
Vn	2.51	3.01	3.04
VIp	5.71	5.17	5.52
VIIp	6.95	5.82	6.20
VIIIp	11.20	10.84	11.83
VIIIn	7.09	6.96	7.13

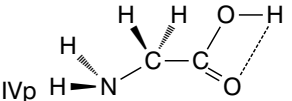
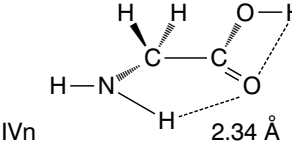
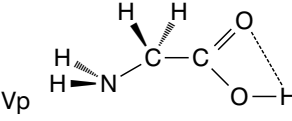
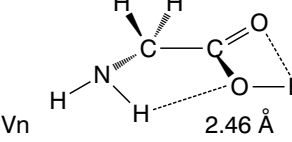
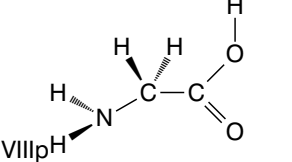
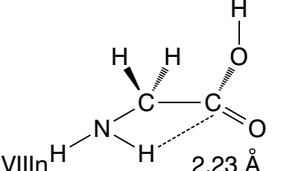
<sup>a</sup> Császár.<sup>25</sup> MP2/6-311++G\*\* calculations.

whereas the original ACM method, proposed by Becke,<sup>34</sup> employs the Perdew91 functional.) They obtained similar results for the energy as those found here. The differences between these two studies can be mainly attributed to basis set. Andzelm et al.<sup>43</sup> showed that, for DFT calculations, basis sets of this size (DZP) developed for Hartree–Fock do not perform as well as those optimized specifically for DFT.<sup>35</sup>

### CORRELATION OF HYDROGEN BONDED INTERACTIONS WITH STABILITY OF GLYCINE CONFORMERS

We observed that the presence of a hydrogen bond can strongly influence the relative stability of the glycine conformers. In Table IX, we compare the relative energies of IVp and IVn, Vp and Vn, and VIIIp and VIIIn. In each of these pairs of structures, the nonplanar form has one additional hydrogen bond and is at least 2 to 3 kcal/mol more stable than the corresponding planar conformation. For example, IVp has one hydrogen bond between the O–H hydrogen atom and the C=O oxygen atom. The MP2 calculations by Császár<sup>24</sup> predict that IVp is 4.7 kcal/mol higher in energy than Ip. Our DFT BLYP and ACM results place IVp at 5.08 and 5.18 kcal/mol, respectively. The relative energy of the corresponding nonplanar structure, IVn, with one additional hydrogen bond of length 2.34 Å between the NH<sub>2</sub> hydrogen atom and the C=O oxygen atom is

**TABLE IX.**  
**Effects of Hydrogen Bonded Interactions on Stability (kcal / mol Relative to Ip) of Glycine Conformers.**

Conformer	MP2 / 6-311 + + G** <sup>a</sup>	BLYP / DZVP	ACM / DZVP
 IVp	4.68	5.08	5.18
 IVn	1.27	1.72	1.59
 Vp	5.38	5.27	5.65
 Vn	2.51	3.01	3.04
 VIIIp	11.20	10.84	11.83
 VIIIIn	7.09	6.96	7.13

<sup>a</sup> Császár.<sup>25</sup> MP2 / 6-311 + + G\*\* calculations.

estimated to be 1.6 kcal/mol by ACM, a decrease of 3.6 kcal/mol from the IVp structure. Similarly, the MP2/6-311 + + G\*\* calculations predict that IVn lies at 1.27 kcal/mol, 3.4 kcal/mol lower than IVp. With the BLYP nonlocal functional, the IVn structure at 1.72 kcal/mol is also about 3.4 kcal/mol lower than IVp.

The planar structure Vp is similar to IVp in that there is a hydrogen bond between the hydroxyl hydrogen and the carbonyl oxygen, which is now trans to the amine nitrogen instead of cis (Table IX). The relative energy of this structure is 5.38 kcal/mol at the MP2/6-311 + + G\*\* level of theory, close to the energy of IVp. We again see that the BLYP and ACM DFT results (5.3 and 5.7 kcal/mol, respectively) are close to the correlated post-HF value and all yield Vp to be slightly

higher in energy than the corresponding structure with the carbonyl cis to the nitrogen (IVp). The related nonplanar structure Vn has one additional, although rather long, hydrogen bond (2.46 Å) between one of the hydrogen atoms of the NH<sub>2</sub> group and the hydroxyl oxygen. Hence, we expect that Vn should be lower in energy than Vp, and find that, indeed, it is roughly 3 kcal/mol lower in energy by both post-HF and density functional methods. This might be somewhat more than would be expected from a long hydrogen bond, but it also replaces a repulsive interaction between the lone pairs on the oxygen and nitrogen atoms.

VIIIp has no internal hydrogen bond because the hydrogen atoms of the NH<sub>2</sub> group point away from the carbonyl oxygen and the carboxylic acid group is in a trans conformation. Of the glycine

structures studied here, VIIIp has the highest energy: it is 11.2 kcal/mol higher than Ip at the MP2 level of theory. The DFT results are within 0.6 kcal/mol of the correlated post-HF calculations, ranging from 10.8 kcal/mol for BLYP to 11.8 kcal/mol for ACM. The nonplanar VIII<sub>n</sub> structure has one hydrogen bond (2.23 Å in length) between the hydrogen of the NH<sub>2</sub> group and the carbonyl oxygen atom. The relative energy of this conformer is calculated to be 7.1 kcal/mol higher with MP2 and 7.1 kcal/mol and 7.0 kcal/mol with ACM and BLYP, respectively. Again, it is some 3–4 kcal/mol less than the non-hydrogen-bonded structure.

The slightly greater sensitivity of the relative energies of VIIIp (and VIIp) to theoretical method can perhaps be rationalized in terms of a decreased opportunity for cancellation of errors in the description of the hydrogen bonds and other weak, nonbonded, interactions. These latter structures have no hydrogen bonds. Absolute errors in describing the hydrogen bonds and weak interactions, which are probably in the range of several tenths of a kilocalorie/mole for current functionals, tend to cancel when conformers involving hydrogen bonds are compared with Ip, but they show up when VIIIp (or VIIp) is considered.

Thus, in conclusion, we see that all methods, both post-HF and density functional, indicate that the presence of a hydrogen bond together with removal of unfavorable N···O interactions in the *nonplanar* structures reduces their energy by about 3.0 kcal/mol relative to the planar conformation.

#### SENSITIVITY OF O4—H10···N1 HYDROGEN BOND DISTANCE IN IIp TO CHOICE OF BASIS SET AND THEORETICAL METHOD

IIp and II<sub>n</sub> are the only two glycine structures with a hydroxyl hydrogen donor–nitrogen acceptor type of hydrogen bond (Fig. 2). In Table X we show the dramatic variation of the O4—H10···N1 hydrogen bond distance in IIp as a function of theoretical method. As a point of comparison, we also tabulate the bifurcated N1—H···O5 hydrogen bond distance in Ip calculated at various levels of theory. We observe that the N1—H···O5 distances calculated with the post-HF methods (CCSD and MP2) are slightly shorter (between 0.05 Å and 0.07 Å) than with the BP86, PP86, BP91, and BLYP nonlocal functionals. The ACM method predicts a N1—H···O5 distance of 2.80 Å which, although still longer, is within 0.02 and 0.03 Å of the CCSD and MP2 values, respectively. On the other hand, the BP86, PP86, and BP91 nonlocal

**TABLE X.**  
Variations of N<sub>1</sub>—H···O<sub>5</sub> Distance in Ip and O<sub>4</sub>—H<sub>10</sub>···N<sub>1</sub> Distance in IIp As a Function of Basis Set and Theoretical Method.

Method	N <sub>1</sub> —H···O <sub>5</sub> distance Ip	O <sub>4</sub> —H <sub>10</sub> ···N <sub>1</sub> distance IIp	Δ <i>E</i> <sub>IIp–Ip</sub>
CCSD / DZP <sup>a</sup>	2.78	1.91	1.43
MP2 / TZ2P	2.77	1.87	0.40
BP86 / DZVP	2.84	1.85	−0.67
PP86 / DZVP	2.83	1.85	−0.25
BP91 / DZVP	2.82	1.84	−0.49
BLYP / DZVP	2.83	1.90	0.17
ACM / DZVP	2.80	1.87	0.11

<sup>a</sup> Hu et al.<sup>28</sup>

functionals give O4—H10···N1 hydrogen bond distances shorter than those predicted by MP2, CCSD, BLYP, and ACM by as much as 0.07 Å.

The shortening of the O4—H10···N1 in IIp and the lengthening of the N1—H···O5 in Ip could result in a stabilization of IIp relative to Ip in BP86, PP86, and BP91. We would expect in general that, as the hydrogen bond length decreases, the energy becomes more favorable. Of course, there are other subtle and competing effects (such as steric repulsion and lone-pair/lone-pair interaction) that determine the conformational stability, but it appears that the correct description of the hydrogen bonded interaction is crucial to predict the relative energies between IIp/II<sub>n</sub> and Ip. It is interesting to note in this connection that the ACM method derived by Becke,<sup>34</sup> which gives good agreement with the large basis set MP2 results, was not specifically optimized for hydrogen bonded or weakly interacting systems. In any event, it appears that the conformations of glycine which contain this variety of different H-bonded and nonbonded interactions should provide a very stringent test for new functionals and classical force field representations for molecular mechanics as they are developed.

#### NATURE OF GLYCINE STATIONARY POINTS CALCULATED USING DFT

To characterize the stationary points calculated with the ACM functional, we calculated the harmonic frequencies of all 13 glycine conformers. In addition, we calculated the vibrational energies from these frequencies which are tabulated in Table XI. Because there are conflicting results regarding

**TABLE XI.**  
**Zero-Point Vibrational Energies (ZPE, kcal / mol) of Glycine Conformers (Number of Imaginary Frequencies Given in Parentheses).**

Conformer	MP2 / 6-311 + + G**		ACM / DZVD		BLYP / DZVD	
	ZPE <sup>b</sup>	ΔZPE <sup>c</sup>	ZPE <sup>b</sup>	ΔZPE <sup>c</sup>	ZPE <sup>b</sup>	ΔZPE <sup>c</sup>
Ip	50.67	0.00	50.62	0.00	48.70	0.00
IIp	50.88 <sup>d</sup> (1)	0.21	50.45 <sup>d</sup> (1)	−0.17	48.87 <sup>d</sup> (1)	0.17
IIIn	51.07	0.40	50.84	0.22	48.35	−0.35
IIIp	51.14	0.47	50.75	0.13	48.82	0.12
IIIn			50.36 <sup>d</sup> (1)	−0.26	48.57 <sup>d</sup> (1)	−0.13
IVp			49.86 <sup>d</sup> (1)	−0.76		
IVn	50.72	0.05	50.35	−0.27		
Vp			50.00 (2)	−0.62		
Vn	50.79	0.12	50.65	0.03		
VIp	50.40	−0.27	50.20	−0.42		
VIIp	50.33	−0.34	49.88	−0.74		
VIIIp			49.44 (2)	−1.18		
VIIIn	50.51	−0.16	50.13	−0.49		

<sup>a</sup> Császár.<sup>25</sup>  
<sup>b</sup> Zero-point vibrational energy (kcal / mol).  
<sup>c</sup> Zero-point vibrational energy relative to Ip (kcal / mol).  
<sup>d</sup> Transition-state structures.

the nature of the IIp, IIIn, IIIp, and IIIIn structures at the HF level,<sup>25</sup> we also calculated the frequencies of these four structures at the BLYP/DZVD level of theory. As shown by Ramek et al.,<sup>25</sup> these four glycine conformers can either be a minimum or a transition state, depending on the choice of basis set (Table III). A more consistent picture emerges from the correlated calculations: BLYP, ACM, and MP2 predict that IIp and IIIIn are transition states, whereas IIIn and IIIp are minima. For the remaining glycine conformers, ACM and MP2 are in agreement predicting that Ip, IVn, Vn, VIp, VIIp, and VIIIn are minima, whereas IVp, Vp, and VIIIp are stationary points with at least one imaginary frequency.

As seen in Table XI, there is general agreement between the zero-point vibrational energies calculated with ACM and the MP2 values (although differences of 0.2 to 0.3 kcal/mol in relative energies occur). The BLYP functional yields zero-point energies which are lower than these by about 2.0 kcal/mol. We note that this latter trend was also observed by Johnson et al.<sup>37</sup> in their assessment of the performance of DFT functionals. Both the ACM and MP2 methods indicate that the zero-point energies can contribute as much as 0.5 kcal/mol to the relative stabilities of the low-energy conformers and even more for the higher energy structures.

### REDERIVATION OF RELATIONSHIP BETWEEN MICROWAVE INTENSITIES AND ENERGIES

The major objective of both the experimental and theoretical studies is to evaluate the relative energetics of the glycine conformers. In the seminal work of Suenram and Lovas,<sup>19</sup> the approximation was made that the vibrational partition functions of the conformers were the same. Now that we are in a position to evaluate this contribution it appears to significantly affect the relative stabilities of the conformers. It also appears that the energy expression derived from the van Vleck–Weisskopf equation<sup>44</sup> derived in this work is actually a free energy and needs to be corrected to make a comparison with the quantum mechanically derived potential energy. Given the importance of these properties, we think it is worthwhile to rederive the relationships between the microwave intensity and the thermodynamics of the conformational equilibrium. The van Vleck–Weisskopf equation<sup>44</sup> expresses the intensity,  $\gamma$ , as:

$$\gamma = \frac{8\pi^2 N f |\mu_{ij}|^2 v^2 \Delta v}{3ckT \left[ (v - v_0)^2 + (\Delta v)^2 \right]} \tag{3}$$

where:  $N$  = the number of molecules/cubic centimeter in the absorption cell;  $f$  = the fraction of



these molecules in the lower of the two states involved in the transition;  $|\mu_{ij}|^2$  = square of the dipole moment matrix element for the transition, summed over the three perpendicular directions in space;  $\nu$  = frequency;  $\nu_0$  = resonant frequency;  $\Delta\nu$  = half-width of the line at half-maximum;  $c$  = velocity of light;  $k$  = Boltzmann constant; and  $T$  = absolute temperature.

Because the peak absorption of the line ( $\gamma^{\max}$ ) occurs very close to the resonant frequency, eq. (3) may be written as:

$$\gamma^{\max} = \frac{8\pi^2 N f |\mu_{ij}|^2 \nu^2}{3ckT \Delta\nu} \quad (4)$$

In the microwave experiment, the rotational transition corresponds to a change in the rotational energy level in a given vibrational and conformational state. The fraction of molecules in a given conformation occupying a particular rotational and vibrational state is given by:

$$f = f_R f_V f_C \quad (5)$$

where  $f_R$ ,  $f_V$ , and  $f_C$  denote the fraction of molecules found in a particular rotational and vibrational and conformational state, respectively.

Substituting this expression for the fraction of molecules in the state of interest [eq. (5)] into eq. (4), we obtain:

$$\gamma^{\max} = \frac{8\pi^2 N f_R f_V f_C |\mu_{ij}|^2 \nu^2}{3ckT \Delta\nu} \quad (6)$$

From this we may obtain the relative intensity of the microwave transition in two conformers (I and II) in terms of their relative populations as:

$$\frac{\gamma_{II}^{\max}}{\gamma_I^{\max}} = \frac{f_{R_{II}} f_{V_{II}} f_{C_{II}} |\mu_{ij}|_{II}^2 (\nu_{II}^2 / \Delta\nu_{II})}{f_{R_I} f_{V_I} f_{C_I} |\mu_{ij}|_I^2 (\nu_I^2 / \Delta\nu_I)} \quad (7)$$

Ultimately, we are interested in the relative thermodynamics between the two conformers. To achieve this we first rearrange eq. (7) to give the relative fraction of molecules in each conformation in terms of the spectroscopic variables and fraction of molecule in the given rotational and vibrational states:

$$\frac{f_{C_{II}}}{f_{C_I}} = \frac{\gamma_{II}^{\max} f_{R_I} f_{V_I} |\mu_{ij}|_I^2 (\nu_I^2 / \Delta\nu_I)}{\gamma_I^{\max} f_{R_{II}} f_{V_{II}} |\mu_{ij}|_{II}^2 (\nu_{II}^2 / \Delta\nu_{II})} \quad (8)$$

We may rewrite this equation as:

$$\frac{f_{C_{II}}}{f_{C_I}} = \alpha \frac{f_{R_I} f_{V_I}}{f_{R_{II}} f_{V_{II}}} \quad (9)$$

where the spectroscopic observables and dipole moments have been collected in  $\alpha$ :

$$\alpha = \frac{\gamma_{II}^{\max} |\mu_{ij}|_I^2 (\nu_I^2 / \Delta\nu_I)}{\gamma_I^{\max} |\mu_{ij}|_{II}^2 (\nu_{II}^2 / \Delta\nu_{II})} \quad (10)$$

The free energy difference between the two conformers,  $\Delta G$ , is related to the conformational equilibrium constant or ratio of the populations by:

$$\Delta G = -RT \ln \frac{f_{C_{II}}}{f_{C_I}} \quad (11)$$

or substituting for  $f_{C_{II}}/f_{C_I}$  from eq. (9), we obtain:

$$\Delta G = -RT \ln \alpha \frac{f_{R_I} f_{V_I}}{f_{R_{II}} f_{V_{II}}} \quad (12)$$

Thus, eq. (12) gives us the relationship between the thermodynamics of the conformational equilibria (free energy) and the microwave observations. We note that it also contains the fraction of the population in the rotational and vibrational states involved in the transition. In the microwave experiment carried out by Suenram and Lovas<sup>19</sup> the rotational transition between the 13<sub>5,8</sub>—12<sub>5,7</sub> rotational states in the ground vibrational states was measured for each conformer. To determine the free energy difference between the two conformers, we need to calculate the fraction of each conformer in the particular rotational and ground vibrational states.

The fraction of molecules in a particular rotational state is given as:

$$f_{R_i} = \frac{e^{(-E_{R_i}/kT)}}{Z_{R_i}} = \frac{(2J+1)e^{(-E_{R_i}/kT)}}{\sqrt{(\pi/ABC)_i}(kT/h)^3} \quad (13)$$

where  $E_{R_i}$  is the energy of the rotational state,  $Z_{R_i}$  is the rotational partition function; and  $A$ ,  $B$ , and  $C$  are the principal moments of inertia for the given conformer. Because the rotational transitions are the same in both conformers, the ratio of  $f_{R_I}/f_{R_{II}}$

becomes:

$$\begin{aligned} \frac{f_{R_I}}{f_{R_{II}}} &= \frac{(2J+1)e^{(-E_{R_I}/kT)}}{\sqrt{(\pi/ABC)_I(kT/h)^3}} \\ &= \frac{(2J+1)e^{(-E_{R_{II}}/kT)}}{\sqrt{(\pi/ABC)_{II}(kT/h)^3}} \\ &= \frac{\sqrt{(ABC)_I}e^{(-E_{R_I}/kT)}}{\sqrt{(ABC)_{II}}e^{(-E_{R_{II}}/kT)}} \end{aligned} \quad (14)$$

Thus, substituting in eq. (12), we obtain the desired expression for the conformational free energy difference between the conformers:

$$\Delta G = -RT \ln \left[ \alpha \frac{\sqrt{(ABC)_I} e^{-E_{R_I}/kT} f_{V_I}}{\sqrt{(ABC)_{II}} e^{-E_{R_{II}}/kT} f_{V_{II}}} \right] \quad (15)$$

We note that if we assume that the vibrational partition functions and therefore the  $f_V$  values are equal for the two conformers, this reduces to the identical expression obtained previously (Ref. 19). However, this was presented as the potential energy difference between the two conformers rather than the free energy difference. This followed from the assumption by these investigators that the energy difference,  $\Delta E$ , rather than the free energy difference, was given by the right-hand side of eq. (11). Basically, this assumes that both the rotational and vibrational partition functions are the same and one loses the contribution from the rotational free energy as well.

We may obtain the expression for the conformational energy difference,  $\Delta E$ , by expressing the fraction of molecules in a given conformation in terms of the sum over all states of the particular conformation divided by the total partition function. That is:

$$f_{C_i} = \frac{e^{-E_{C_i}/kT} \sum e^{-(E_V/kT)}_i \sum e^{-(E_R/kT)}_i}{Z} \quad (16)$$

where the numerator contains the exponential of the potential energy of conformer  $i$ , multiplied by the sum over all vibrational states of the conformer, and over all rotational states of this conformer. The total partition function contains analogous sums over all conformers in the system. The ratio of the population of the two conformers I and

II is then given by:

$$\frac{f_{C_{II}}}{f_{C_I}} = \frac{e^{-E_{C_{II}}/kT} Z_{V_{II}} Z_{R_{II}}}{e^{-E_{C_I}/kT} Z_{V_I} Z_{R_I}} \quad (17)$$

where we have noted that the sum over all vibrational and rotational states of the Boltzmann factor is the corresponding partition function,  $Z$ , for the particular conformational state. (It is clear, from this equation, that the assumption that  $F_{C_{II}}/F_{C_I} = e^{-\Delta E/kT}$  implies that both the vibrational and rotational partition functions are the same in both states.) Substituting eq. (17) into eq. (9):

$$\frac{f_{C_{II}}}{f_{C_I}} = \alpha \frac{f_{R_I} f_{V_I}}{f_{R_{II}} f_{V_{II}}} \quad (9)$$

we obtain:

$$e^{((E_{C_I} - E_{C_{II}})/kT)} \frac{Z_{V_{II}} Z_{R_{II}}}{Z_{V_I} Z_{R_I}} = \alpha \frac{f_{R_I} f_{V_I}}{f_{R_{II}} f_{V_{II}}} \quad (18)$$

as given in (13), the fraction of molecules in a particular rotational state in conformer  $i$  is:

$$f_{R_i} = \frac{e^{(-E_{R_i}/kT)}}{Z_{R_i}} \quad (19)$$

and, similarly, the fraction of molecules in a particular vibrational state in conformer  $i$  is:

$$f_{V_i} = \frac{e^{(-E_{V_i}/kT)}}{Z_{V_i}} \quad (20)$$

Substituting these expressions, eqs. (19) and (20), for the fraction of molecules in a given state into eq. (18) we obtain:

$$e^{(-\Delta E/kT)} \frac{\alpha [e^{(-E_{R_I}/kT)} e^{(-E_{V_I}/kT)}]}{e^{(-E_{R_{II}}/kT)} e^{(-E_{V_{II}}/kT)}} \quad (21)$$

or:

$$\Delta E = kT \ln \left\{ \frac{\alpha [e^{(-E_{R_I}/kT)} e^{(-E_{V_I}/kT)}]}{e^{(-E_{R_{II}}/kT)} e^{(-E_{V_{II}}/kT)}} \right\} \quad (22)$$

Thus, eqs. (15) and (22) give us the energy and free energy differences between the two conformers we seek in terms of the experimental microwave observables and the vibrational energies.

## REDETERMINATION OF ENERGY DIFFERENCE BETWEEN IP AND IIN

Next, we compare our expressions for the energies and the one used by Suenram and Lovas.<sup>19</sup>

Also, we re-evaluate the experimental relative energies between Ip and IIn using the above-derived expressions for free energy, conformational energy, and ground state energy. As can be observed in Table XII, the expression for  $W_c$  used by Suenram and Lovas<sup>19</sup> and our expression for  $\Delta G$  are identical. In their study, because they assumed that the  $f_v$  terms are the same for both Ip and IIn, they obtained a value of 1.4 kcal/mol for  $W_c$ . If we used the vibrational frequencies from the MP2 study of Császár,<sup>24</sup> we arrived at 2.1 kcal/mol as the free energy difference between Ip and IIn. This experimental value is not too far from the estimate of 1.7 kcal/mol given in the electron diffraction study of Iijima et al.,<sup>23</sup> which assumed that two quasi-degenerate conformers (IIn and IIIp) are populated. Thus, it appears that both the microwave and electron diffraction studies are in reasonable agreement regarding the free energy difference between Ip and IIn.

Furthermore, from Table XII, we can see that the expression for the ground state ( $\Delta E_0$ ) and conformational ( $\Delta E_c$ ) energies differ from the free energy by the terms involving the rotational free energy, which is proportional to the square root of the product of the moments of inertia, and the vibrational free energy ( $f_v$ 's). If we utilized the measured values for  $\gamma_{\max}$ ,  $\nu$ ,  $\Delta\nu$ ,  $\mu$ , and  $E_R$  from Suenram and Lovas,<sup>19</sup> we obtain 1.37 kcal/mol as the rederived experimental value for  $\Delta E_0$ . Thus,

we see that by assuming that the  $f_v$ 's are equal in both Ip and IIn, and by using the wrong definition for  $W_c$  Suenram and Lovas fortuitously arrived at the same answer.

If we use the vibrational frequencies for Ip and IIn obtained in the MP2 study of Császár<sup>24</sup> to calculate the zero-point vibrational energies for Ip and IIn and the value for  $\Delta E_0$ , we obtain 1.0 kcal/mol as the rederived experimental value for the conformational energy difference ( $\Delta E_c$ ). Thus, the effect of the vibrational energies is to further reduce  $\Delta E_0$  by ca. 0.4 kcal/mol. Use of frequencies calculated from HF/DZP, ACM/DZVD or ACM/TZVP, or MP2 yields experimental values for  $\Delta E_c$  in the range of 0.97 to 1.14 kcal/mol. We, therefore, suggest that an uncertainty in the experimental  $\Delta E_c$  of  $\pm 0.1$  kcal/mol should be associated with the lack of precise knowledge of the vibrational frequencies. Adding this to the previous uncertainty associated with the dipole moments and with the sensitivity of the spectrometer yields an error of  $\pm 0.5$  kcal/mol to the energy of 1 kcal/mol. Thus, we are seeing that a consistent picture has emerged regarding the relative electronic energy difference between Ip and IIn. The ACM and BLYP relative energies between IIn and Ip of 0.12 and 0.08 kcal/mol, respectively, are somewhat lower than the range of the revised experimental energy. The MP2 value lies at the lower end of the revised experimental error bars. The highest quality post-HF results [CCSD(T)/DZP] of 1.0 kcal/mol lie near the center of the range 0.5 to 1.5 kcal/mol. In Table XIII, we tabulate the results for the conformational energies and ground state energies for all 13 glycine conformers.

Also in Table XIII, we present results for the relative free energies. Although Ip and IIn are predicted by DFT to be almost isoenergetic, *entropic contributions significantly effect the relative stabilities*. The free energy of IIn is calculated to be 1.63 and 1.52 kcal/mol above Ip by the MP2 and ACM methods, respectively. These are close to the experimental value, being on the low side due to the energetic contribution. In fact, inspection of Table XII reveals that the MP2 and ACM vibrational contributions to the free energy, and  $\Delta ZPE$  can each differ by several tenths of a kilocalorie/mole. The differences in  $\Delta G_{473}$  as calculated by MP2 and ACM range from 0.1 kcal/mol (IIn) to 1.8 kcal/mol (VIIp). Firm values of the vibrational frequencies would be required before comparisons can be made beyond the level of several tenths of a kilocalorie/mole.

**TABLE XII.**  
**Comparison of Relationship Between Free Energy, Conformational Energy, and Ground State Energy with Microwave Observables.**

Energy expression	Experimental energy (kcal/mol)
$\Delta G = -kT \ln \left[ \alpha \frac{\sqrt{(ABC)_I} e^{-E_{R_I}/kT} f_{V_I}}{\sqrt{(ABC)_{II}} e^{-E_{R_{II}}/kT} f_{V_{II}}} \right]$	$\Delta G = 2.1$
$W_c = -kT \ln \left[ \alpha \frac{\sqrt{(ABC)_I} e^{-E_{R_I}/kT} f_{V_I}}{\sqrt{(ABC)_{II}} e^{-E_{R_{II}}/kT} f_{V_{II}}} \right]$	$W_c = 1.4^a$
$\Delta E_0 = -kT \ln \left[ \alpha \frac{e^{-E_{R_I}/kT}}{e^{-E_{R_{II}}/kT}} \right]$	$\Delta E_0 = 1.37$
$\Delta E_c = -kT \ln \left[ \alpha \frac{e^{-E_{R_I}/kT}}{e^{-E_{R_{II}}/kT}} \right] - \Delta ZPE$	$\Delta E_c = 1.0$

<sup>a</sup> Suenram and Lovas.<sup>20</sup>

**TABLE XIII.** Computed Conformational Energies ( $E_C$ ), Zero-Point Vibrational Energies (ZPE), Ground-State Energies ( $E_0$ ), and Total Free Energies at 473 K ( $G_{473}$ ) Due Only to Nuclear Motion (Rotation, Translation, and Vibration) for All Known Glycine Conformers.

Conf.	Exp <sup>a</sup>			MP2 / 6-311 + + G** <sup>b</sup>				ACM / DZVP <sup>c</sup>			
	$\Delta E_C$	$\Delta E_0$	$\Delta G$	$\Delta E_C$	$\Delta ZPE$	$\Delta E_0$	$\Delta G_{473}^d$	$\Delta E_C$	$\Delta ZPE$	$\Delta E_0$	$\Delta G_{473}^d$
Ip	0.0		0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IIp				0.57	0.21	0.78	3.36	0.11	−0.17	−0.06	1.96
IIIn	1.0	1.4	2.1 (1.7 <sup>e</sup> )	0.51	0.40	0.91	1.63	0.08	0.23	0.31	1.52
IIIp				1.59	0.47 <sup>f</sup>	2.06 <sup>f</sup>	2.87 <sup>f</sup>	1.56	0.13	1.69	1.72
IIIIn				1.45				1.59	−0.25	1.34	3.14
IVp				4.68				5.18	−0.76	4.42	5.02
IVn				1.27	0.05	1.32	1.85	1.59	−0.11	1.48	1.59
Vp				5.38				5.65	−0.61	5.04	7.59
Vn				2.51	0.12	2.63	3.05	3.04	0.03	3.07	2.76
VIp				5.71	−0.27	5.54	5.37	5.52	−0.41	5.11	4.44
VIIp				6.95	−0.34	6.61	6.67	6.20	−0.73	5.47	4.90
VIIIp				11.20				11.83	−1.17	10.66	13.27
VIIIIn				7.09	−0.16	6.93	7.40	7.13	−0.49	6.64	6.32

<sup>a</sup> Revised Suenram and Lovas<sup>19</sup> results to account for the difference in the fractional occupancies of the absorbing state for conformers Ip and IIIn.  
<sup>b</sup> Császár.<sup>24</sup>  
<sup>c</sup> In the ACM calculations, the ACM / DZVD frequencies are used.  
<sup>d</sup> A factor of 1 / 2 RT is added to the free energy of the transition state to account for the missing degree of freedom.  
<sup>e</sup> Iijima et al.<sup>22</sup> The value assumes that two quasidegenerate states are populated.  
<sup>f</sup> Present work.

COMPOSITION OF GLYCINE PRESENT IN THE PHASE

After having calculated the relative free energies, we can now turn our attention to analyzing the composition of glycine in the gas phase. The Boltzmann equilibrium distribution ( $f_C$ ) of glycine among the eight minima is obtained from the free energy values calculated with MP2 and ACM using eq. (6). Results are tabulated in Table XIV along with the relative mole fractions obtained from the electron diffraction study<sup>22</sup> for comparison.

The Császár MP2 study<sup>24</sup> predicted that conformer Ip (71.1%) would be the major component of glycine in the gas phase at 473 K. The concentrations of IIp and IVn were estimated by MP2 to be 12.5% and 9.9%, respectively. Of the remainder, ~6% of the vapor is composed of conformer IIIp (3.4%) and Vn (2.8%), and the remaining few tenths of a percent contains VIp, VIIp, and VIIIIn. ACM estimates that 62.1% of glycine is in conformer Ip and that IIIn, IIIp, and IVn are almost equally populated at 12.3%, 10.0%, and 11.4%, respectively. Vn makes up about 3.3% of the gas, whereas the remaining 1% is composed of VIp, VIIp, and VIIIIn. Hence, these two correlated methods, MP2

**TABLE XIV.** Calculated Mole Fraction ( $f_C$  in Percent), of Glycine Conformers in Gas Phase at 473 K.

Conf.	Exp.	HF / 6-31G* <sup>a</sup>	MP2 / 6-311 + + G** <sup>b</sup>	ACM / DZVP
Ip	76 <sup>c</sup>	45	71.1	62.1
IIIn	24 <sup>c</sup>	7.5	12.5	12.3
IIIp		35.0	3.4	10.0
IVn		7.5	9.9	11.4
Vn		5.0	2.8	3.3
VIp		0.0	0.2	0.6
VIIp		0.0	0.1	0.3
VIIIIn		0.0	0.0	0.1

<sup>a</sup> Jensen and Gordon study.<sup>23</sup>  
<sup>b</sup> Császár's MP2 study.<sup>24</sup>  
<sup>c</sup> Electron diffraction results.<sup>22</sup>  
<sup>d</sup> Suenram and Lovas.<sup>19</sup>

and ACM, yield reasonably consistent results. However, they give a very different composition than the HF study of Jensen and Gordon.<sup>23</sup> HF predicted that 45% of glycine would be in Ip, whereas the second most dominant conformer would be IIIp at 35%. The rest would be composed of IIIn (7.5%), IVn (7.5%), and Vn (5%).

The principal differences in the relative populations among the theoretical methods can be traced to the sensitivity of the lowest vibrational mode of IIIp. At the HF level of theory, this lowest C<sub>2</sub>—C<sub>3</sub> torsional mode has a frequency of only 16 cm<sup>-1</sup>. MP2 and ACM predict a vibrational frequency of 86 and 69 cm<sup>-1</sup>, respectively. (The B3LYP calculation of Barone et al.<sup>42</sup> estimates that the lowest frequency of IIIp is 47 cm<sup>-1</sup>.) Thus, we observe a 70 cm<sup>-1</sup> variation for this mode. We do not observe such a large spread in the calculated frequencies for other conformers. For example, the lowest mode in Ip which also corresponds to the C<sub>2</sub>—C<sub>3</sub> torsional mode is calculated to be 72 cm<sup>-1</sup> at the HF level of theory. MP2, ACM, and B3LYP estimate a frequency of 54, 88, and 75 cm<sup>-1</sup>, respectively. To demonstrate the sensitivity in calculating the free energies and hence the equilibrium distributions, for example, we used the MP2 energies and frequencies and arbitrarily changed the lowest frequency in IIIp from 86 cm<sup>-1</sup> to 16 cm<sup>-1</sup>. This changed the IIIp population from 3.4% to 16.0%. This reflects the necessity to accurately calculate the vibrational frequencies, especially low frequency modes, to account for the free energy.

Hence, the picture that emerges from the theoretical calculations of glycine, as well as experimental studies, is that of a complex equilibrium, more complex than had previously been accounted for by experiment, although it was suspected. Both theoretical methods, post-Hartree-Fock and density functional theory, indicate that there are at least four significant components in the gas phase: structure Ip, which is the unanimous choice of all methods to be the most stable; IIn; IIIp; and IVn. This complex equilibrium explains in part why the experimental techniques, whose resolution limited them to two state models, were unable to come to a definitive result with respect to the structure of glycine in the gas phase. As noted, both theoretical and experimental methods all agree that the most stable and most prevalent conformer is structure Ip. The correlated theoretical methods also agree in the determination that conformers IIn and IVn (IVn is a conformer previously not considered in the experimental work), both exist in the gas phase at concentrations of approximately 10%. Conformer IIIp, which was assumed to exist in some of the experimental studies, is also found to be significant and density functional theory also finds it to be over 10%, whereas the post-Hartree-Fock method yields a concentration of only 3.4%. This

difference, as discussed in the present work, is due to the sensitivity of this concentration to a single low frequency mode which was also responsible for Hartree-Fock finding it to be present at 35%. In addition to these, there are still other conformers which are present in minor and less significant concentrations including Vn, which is found to make up approximately 3% of the mixture at 473 K, as well as VIp, VIIp, and VIIIIn, which account for fractions of a percent. The common picture that emerges from the two different methods, density functional and post-Hartree-Fock, is encouraging in the fact that, although it expands the experimental interpretation, it does not contradict it. Thus, we expect that we are achieving a reasonable picture of the composition of this important molecule in the gas phase. Clearly, higher level and more accurate calculations, in which geometry relaxation and vibrational free energy are taken into account, would be useful. This molecule should serve as an important test system for theoretical methods, both quantum and classical (molecular mechanical), to describe larger systems of biological interest.

---

## Conclusions

Here we summarize the important results and conclusions of this study with regard both to the accuracy and applicability of density functional methods and the energetic surface of glycine that results from these calculations, as well as post-Hartree-Fock calculations and experiment. First, from the studies carried out here, it appears that two of the more recent and widely used density functional methods—the BLYP nonlocal functional and the ACM method—are able to account reasonably for the energetics of glycine and the relative energies of the two lowest conformations, Ip and IIn, although the energy differences calculated by these methods are on the low side. In general, the relative energies obtained from density functional theory are in good agreement with those obtained from MP2 and specifically are within 0.5 kcal/mol of the MP2 results for conformers within 6 kcal/mol of the lowest energy conformer and within 1 kcal/mol for more strained configurations.

The relative energies of glycine that emerge from the combination of density functional, post-Hartree-Fock, and experimental studies suggest

that structure Ip is the global minimum and that IIn is roughly 0.5 to 1.0 kcal/mol higher in energy. The experimental energy difference has been revised here, through the use of the vibrational frequencies, to account for the vibrational free energy. The effect of the vibrational free energy is to reduce the energy difference from approximately 1.4 kcal/mol to 1 kcal/mol  $\pm$  0.5. Both MP2 and ACM predict that IIIp is slightly higher in energy, roughly 1.5 kcal/mol higher than Ip. A structure not previously considered, IVn, has been found by the post-Hartree-Fock results to be only 1.3 kcal/mol above Ip, lower in energy than IIIp, and this is qualitatively confirmed by ACM which finds the structure at 1.7 kcal/mol, slightly higher than structure III in this case.

Consideration of zero-point and thermal vibrational contributions is essential for a quantitative interpretation. These energies can be greater than 1.0 kcal/mol. Uncertainties for the vibrational contributions are several tenths of a kilocalorie/mole, using the best available quantum chemical techniques.

Finally, in the gas phase at 473 K, glycine is composed of a number of conformations that have similar free energies (within 3.0 kcal/mol) making experimental interpretation difficult. Ip is the main conformer ( $\sim$  60–70%) present in the gas phase at 473 K. The remainder of the gas is composed of IIn, IIIp, IVn, and Vn. The free energies and, hence the relative populations, are very sensitive to the vibrational frequencies.

## Acknowledgments

The authors thank Drs. Carl Ewig and Marvin Waldman for their help and interest in understanding the relationship between microwave intensity and energy.

## References

1. A. D. Becke, *Phys. Rev.*, **A33**, 2786 (1988).
2. C. Lee, W. Yang, and R. G. Parr, *Phys. Rev.*, **B37**, 785 (1988).
3. J. P. Perdew, *Electronic Structures of Solids*, Akademie Verlag, Berlin, 1991.
4. F. Sim, A. St-Amant, I. Papai, and D. R. Salahub, *J. Am. Chem. Soc.*, **114**, 4391 (1992).
5. W. Kohn and L. J. Sham, *Phys. Rev.*, **A140**, 1133 (1965).
6. J. Bajorath, D. H. Kitson, G. Fitzgerald, J. Andzelm, J. Kraut, and A. T. Hagler, *Proteins*, **9**, 217 (1991).
7. J. Bajorath, J. Kraut, Z. Li, D. H. Kitson, and A. T. Hagler, *Proc. Natl. Acad. Sci. USA*, **88**, 6423 (1991).
8. J. Bajorath, Z. Li, G. Fitzgerald, D. H. Kitson, M. Farnum, R. M. Fine, J. Kraut, and A. T. Hagler, *Proteins*, **11**, 263 (1991).
9. A. T. Hagler, M.-J. Hwang, and R. Sharon, *J. Am. Chem. Soc.*, **118**, 3759 (1996).
10. C. S. Ewig, C. Liang, P. L. Polavarpu, T. R. Stouch, and A. T. Hagler, unpublished results.
11. H. Sun, S. J. Mumby, J. R. Maple, and A. T. Hagler, *J. Phys. Chem.*, **99**, 5873 (1995).
12. C. X. Liang, C. S. Ewig, T. R. Stouch, and A. T. Hagler, *J. Am. Chem. Soc.*, **116**, 3904 (1994).
13. A. T. Hagler, and C. S. Ewig, *Comput. Phys. Commun.*, **84**, 131 (1994).
14. J. R. Maple, M. J. Hwang, T. P. Stockfish, U. Dinur, and A. T. Hagler, *J. Comput. Chem.*, **15**, 162 (1994).
15. S. Vishveshwara and J. A. Pople, *J. Am. Chem. Soc.*, **99**, 2422 (1977).
16. R. D. Brown, P. D. Godfrey, J. W. Storey, and M. P. Bassez, *JCS Chem. Commun.* 547 (1978).
17. R. D. Suenram and F. J. Lovas, *J. Mol. Spec.*, **72**, 372 (1978).
18. H. L. Sellers and L. Schäfer, *J. Am. Chem. Soc.*, **100**, 7728 (1978).
19. R. D. Suenram and F. J. Lovas, *J. Am. Chem. Soc.*, **102**, 7180 (1980).
20. L. Schäfer, H. L. Sellers, F. J. Lovas, and R. D. Suenram, *J. Am. Chem. Soc.*, **102**, 6566 (1980).
21. C. E. Dykstra, R. A. Chiles, and M. D. Garrett, *J. Comput. Chem.*, **2**, 266 (1981).
22. K. Iijima, K. Tanaka, and S. Onuma, *J. Mol. Struct.*, **246**, 257 (1991).
23. J. H. Jensen and M. S. Gordon, *J. Am. Chem. Soc.*, **113**, 7971 (1991).
24. A. G. Császár, *J. Am. Chem. Soc.*, **114**, 9568 (1992).
25. M. Ramek, V. K. W. Cheng, R. F. Frey, S. Q. Newton, and L. Schäfer, *J. Mol. Struct. (Theochem)*, **235**, 1 (1991).
26. R. F. Frey, J. Coffin, S. Q. Newton, M. Ramek, V. K. W. Cheng, F. A. Momany, and L. Schäfer, *J. Am. Chem. Soc.*, **114**, 5369 (1992).
27. C.-H. Hu, M. Shen, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **115**, 2923 (1993).
28. P. D. Godfrey and R. D. Brown, *J. Am. Chem. Soc.*, **117**, 2019 (1995).
29. P. D. Godfrey, R. D. Brown, and F. M. Rodgers, *J. Mol. Struct.*, **376**, 65 (1996).
30. A. St-Amant and D. R. Salahub, *Chem. Phys. Lett.*, **169**, 387 (1990).
31. J. Baker, A. C. Scheiner, and J. W. Andzelm, Unpublished results.
32. J. P. Perdew and Y. Wang, *Phys. Rev.*, **B33**, 8822 (1986).
33. J. P. Perdew and X. Wang, *Phys. Rev.*, **B45**, 13244 (1992).
34. A. D. Becke, *J. Chem. Phys.*, **98**, 5648 (1993).
35. N. Godbout, D. R. Salahub, J. W. Andzelm, and E. Wimmer, *Can. J. Chem.*, **70**, 60 (1992).

36. H. B. Schlegel, *Ab Initio Methods in Quantum Chemistry*, Wiley, New York, 1987.
37. B. G. Johnson, P. M. W. Gill, and J. A. Pople, *J. Chem. Phys.*, **98**, 5612 (1993).
38. J. Baker, J. W. Andzelm, A. Scheiner, and B. Delley, *J. Chem. Phys.*, **101**, 8894 (1994).
39. S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Chem.*, **58**, 1200 (1980).
40. A. D. Becke, *Phys. Rev.*, **A38**, 3098 (1988).
41. F. Leij, C. Adamo, and V. Barone, *Chem. Phys. Lett.*, **230**, 189 (1994).
42. V. Barone, C. Adamo, and F. Leij, *J. Chem. Phys.*, **102**, 364 (1995).
43. J. Andzelm, J. Baker, A. Scheiner, and M. Wrinn, *Int. J. Quant. Chem.*, **56**, 733 (1995).
44. C. H. Townes and A. L. Schawlow, *Microwave Spectroscopy*, McGraw-Hill, New York, 1955.